0-315-01213-7

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada K1A 0N4

49071

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

LARRY PELENSKY

Date of Birth — Date de naissance

Sept 22, 1953

Country of Birth — Lieu de naissance

Canada

Permanent Address — Résidence fixe

3931-51 st Edmonton

Title of Thesis — Titre de la thèse

Construct and Predictive Validity of the Aptitude Inventory

University — Université

University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

M. Ed.

Year this degree conferred — Année d'obtention de ce grade

1980

Name of Supervisor — Nom du directeur de thèse

George Fitzsimmons

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date

~~July~~ 22 June 13, 1980

Signature

Larry Pelensky

NL-91 (4/77)

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Ottawa, Canada
K1A 0N4

NL-339 (Rev. 8/80)

THE UNIVERSITY OF ALBERTA

CONSTRUCT AND PREDICTIVE VALIDITY OF THE APTITUDE INVENTORY

BY

© LARRY PELENSKY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1980

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Reserach, for acceptance, a thesis entitled Construct and Predictive Validity of the Aptitude Inventory submitted by Larry Pelensky in partial fulfillment of the requirements for the Degree of Master of Education.

.....
Supervisor

.....

.....

Date ...6...June...1980...

ABSTRACT

This research was an investigation of the construct and predictive validities of the Aptitude Inventory. To study these two concepts, the scales of the Aptitude Inventory, the Edwards Personal Preference Schedule and the California Psychological Inventory, were compared, and the ability of an actuary to predict "success" in real estate sales using these three instruments was examined.

Study A, a correlated sample of 2425 executives assessed by industrial psychologists provided data for correlation matrix and factor analysis of the three instruments. Study B subjects (N=22) were seeking employment with one firm as real estate sales persons. Ten psychological tests, an interview, and production records for first year were obtained for all subjects in Study B.

Results indicate that the Aptitude Inventory does not measure the same dimensions as do scales of the Edwards Personal Preference Schedule and the California Psychological Inventory. Study B results suggest that a combination of weighted scales of either the C.P.I. or Edwards are better predictors of "success" in first year real estate sales than the (Sales) Aptitude Inventory. The best R^2 value was .746 using 2 scales of the C.P.I. and 2 scales of the E.P.P.S. to predict success in real estate sales.

ACKNOWLEDGEMENTS

I would like to thank the following people who were of assistance in completing this research. Of great help was the staff of A. W. Fraser and Associates, especially Al Fraser, Jim Wuest, Pat Mitchell and Carol Miller who gave their time and support. I also appreciate the advice and counsel given by the members of my committee, Dr. George Fitszimmons, Dr. Gene Fox, and Dr. Ken Ward. Most importantly, I would like to thank my wife, Maureen who was not only supporting and understanding, but willing to give unconditionally of her time.

Larry Pelensky

TABLE OF CONTENTS

CHAPTER		PAGE
1	THE PROBLEM	1
2	REVIEW OF THE LITERATURE	3
	Validity	3
	Validity of the Aptitude Inventory	7
	Personality Testing	11
	California Psychological Inventory	12
	Edwards Personal Preference Schedule	15
	Aptitude Inventory	17
	Similarities of these Instruments	18
	Problems with Personality Tests	24
	Prediction	26
	Models of Clinical Judgment	26
	The Validity of Clinical Judgment	27
	Predictive Validity Studies of the Aptitude Inventory	29
	Summary	31
3	PROCEDURE AND DESIGN	33
	Study A: Construct Validity	33
	Data Gathering Procedure	33
	Sample	33
	Analysis Procedure	34
	Study B: Predictive Validity	34
	Data Gathering Procedure	34
	Sample	39

CHAPTER		PAGE
	Analysis Procedure	40
	Limitations	41
4	FINDINGS OF THE STUDY	43
	Study A: Construct Validity	43
	Discussion	45
	Additional Findings Pertinent to Study A	45
	Study B: Predictive Validity	48
	Discussion	56
	Additional Findings Pertinent to Study B	56
5	DISCUSSIONS AND IMPLICATIONS	58
	Study A: Construct Validity	58
	Implications	58
	Study B: Predictive Validity	59
	Implications	61
	Conclusions	61
	Suggestions for Further Research	62
	REFERENCES	64
	APPENDICES	
	Appendix 1: Correlations Among Scales of the Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory	71
	Appendix 2: Clinician's prediction of "Success" based on Aptitude Inventory alone, and 10 test battery	78

LIST OF TABLES

TABLE		PAGE
1	Descriptions, phrases and needs for the Aptitude Inventory, Edwards Personal Preference Schedule and California Personality Inventory Scales	20
2	Significant Correlations Between Scales of the Edwards Personal Preference Schedule and of the California Psychological Inventory	23
3	Production Records for 22 Real Estate Sales Persons .	36
4	Market Survey of Real Estate Prices	37
5	Aptitude Inventory Results for 22 Real Estate Sales Persons	40
6	Correlations among conceptually similar scales of the Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory	44
7	Factor Analysis of the Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory	46
8	Correlations Between the criterion and the Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory	50
9	Correlations, Multiple Regressions and R^2 values to predict the criterion	53
10	Two way Anova to determine significant scales	54

CHAPTER 1

THE PROBLEM

Recruitment and selection are two procedures of organizations that are becoming increasingly difficult and costly. Industrial psychologists and other professionals are concerned with what makes a good candidate for a specific job, and how to identify such individuals scientifically. To assist in the identification of suitable candidates, many established psychological tests are used. Further, some tests have been developed that are designed specifically to determine potential or aptitude for specific jobs.

Research in this area has focused on clinical judgment versus actuarial prediction (Meehl, 1954; Holt, 1958, 1970; Sawyer, 1966; Dawes, 1979). Further research has dealt with testing as an adjunct or sole means of applicant assessment (Heinrichs, 1969; Spitzer & McNamara, 1964; Bray & Moses, 1972). A few tests have a great deal of validity research: the California Psychological Inventory and the Edwards Personal Preference Schedule, whereas others have far less research.

A major area of concern that prompted this study, is that too often psychological tests are devised and used to assist in selection systems, but have insufficient research to demonstrate their validity. One such instrument is the Aptitude Inventory. Though it is widely used by industrial psychologists across North America, there is very limited validity research available. The Aptitude Inventory is published in the same form under three titles, the Management Aptitude Inventory, the Employment Aptitude Inventory, and the Sales Aptitude Inventory, all of which have identical items and norms.

Research on the Aptitude Inventory is limited to four sources: (Denton, 1964; Denton, 1969; Guion, 1963; Ross, 1963). These studies do very little to demonstrate the validity of this instrument (Siegel, 1959; Buros, 1969). It follows therefore that there is a definite need for more validity research of the A.I.

This study is an investigation of the construct and predictive validity of the Aptitude Inventory. The research questions are designed to determine both the validity of the four A.I. constructs, and the validity of this instrument in predicting "success" in the field of sales. One specific interest in this study is the relationship between the scales of the A.I. and the Edwards Personal Preference Schedule and the California Psychological Inventory. Also of major interest, is the relative ability of actuaries using the three above-mentioned instruments to predict success in sales, one of the three aspects it purports to measure.

This study is of value at both a theoretical level, and an operational level. At a theoretical level, rationale for the study focuses on determining the validity of the constructs used in the A.I.. At present, there exists no research to provide evidence that the A.I. does in fact measure the constructs it purports to measure. At an operational level, this study is important because it is concerned with the efficacy of using the A.I. to predict success in the fields of sales. If the A.I. does prove effective as a predictor, it can add significantly to the selection process, and can also make the process more economical.

CHAPTER 2

REVIEW OF THE LITERATURE

At last count, in Tests in Print (Buros, 1974), and the Eighth Mental Measurement Yearbook (Buros, 1978), there were 354 non projective personality tests and questionnaires on the market. While a few of these instruments have generated thousands of validity studies and/or research articles, others, even those in constant use, have had extremely limited research on the subject of their validity. This chapter contains an overview of some of the current ideas regarding the nature of validity. It also reviews the limited validity research of one frequently used personality test, The Aptitude Inventory, and discusses personality testing. Since most psychological tests attempt to improve prediction (clinical judgement), and many validity studies are based on the prediction of a criterion, this chapter ends with a review of the literature regarding clinical judgement. These four considerations provide the theoretical support for this validity study.

Validity

From 1945 to 1955, a plethora of psychological tests were published, with the result that much writing on test validity was done. Historically, three main publications established the background for definitions of validity. They are: Chronbach and Meehl (1955); Loevinger (1957); and Technical Recommendations (1954).

The classical definition of validity, or the definition that has been predominant in psychometric literature is as follows. Validity is the extent to which a test measures what it is supposed to measure (Loevinger, 1957; Chronbach, 1970; Shertser & Linden,

1979). However, as early as 1957, Loevinger argued that classical validity is not a suitable measure of test validity. "The contention of the present monograph is that classical validity is not a suitable basic concept for test theory; it does not provide an adequate basis for test construction" (p. 79).

The American Psychologists Association Committee on psychological tests (1950-1954) stated that validation of psychological tests had not been adequately conceptualized (Chronbach & Meehl, 1955). The Committee later published Technical Recommendations for Psychological Tests and Diagnostic Techniques (1954) and at that time defined four types of validity:

1. Predictive validity
2. Concurrent validity
3. Content validity
4. Construct validity.

Chronbach & Meehl (1955) use this nomenclature, but later reclassify predictive and concurrent validity under one main heading, criterion-oriented validity. They state that criterion-oriented validity measures the extent to which a test can predict a criterion. In studying criterion-oriented validity, the researcher administers a test to a group, obtains an independent criterion measure on these subjects, then computes a correlation. "If the criterion is obtained some time after this test is given, he is studying predictive validity. If the test score and the criterion score are determined at essentially the same time, he is studying concurrent validity" (Chronbach & Meehl, 1955, p. 58).

Content validity is concerned directly with test items. Technical

Recommendations (1954) suggests that content validity is demonstrated by determining how well the content of the test samples the class of situations or subject matter about which conclusions are to be drawn. Therefore, to determine content validity, in a test composed of a number of subscales, the researcher must demonstrate that the individual items do in fact measure what the scales state they measure.

Construct validity is based on the assumption that the test being studied "reflects a particular construct to which are attached certain meanings" (Chronbach & Meehl, 1955, p. 65). Based on this assumption, the researcher must determine the psychological qualities that the test measures (Technical Recommendations, 1954). "Essentially in studies of construct validity we are validating the theory underlying the test" (Technical Recommendations, 1954, p. 14). There are two main questions subsumed under construct validation; to what extent does the test measure whatever is specified by a given construct, and to what extent does that construct embody a valid hypothesis? Therefore, a user of a test must accept the author's theory before he can accept the author's validation (Chronbach & Meehl, 1955).

Using a slightly different nomenclature, Loevinger (1957), describes construct validity as the most important concept with "all of the other kinds of validity as possible supporting evidence" (p. 92). To Loevinger, validity has three aspects; the substantive component, the structural component and the external component.

"The substantive component of validity is the ability of theory to account for the resultant content" (Loevinger, 1957, p. 97). This component closely resembles content validity. The structural component of validity refers to "the extent to which structural relations between

test items parallel the structural relations of other manifestations of the trait being measured" (Loevinger, 1957, p. 97). The external component of validity is Loevinger's classification for what Chronbach & Meehl classify as criterion-oriented validity, and Technical Recommendations calls concurrent and predictive validity. Loevinger (1957), in commenting on this form of validity, states "the problem of external criteria for validity was ably treated" (p. 107) by Chronbach & Meehl (1955) and Technical Recommendations (1954).

Since the early years of psychological testing, many authors (Shertzer & Linden, 1979; Chronbach, 1970; Wiggins, 1973; Edwards, 1970; Kleinmuntz, 1967) have used different nomenclatures for validity. They also have extended the basic principals outlined by Technical Recommendations (1954), Chronbach & Meehl (1955) and Loevinger (1957). However, none of the basic concepts have changed over the past three decades.

An example of a recent nomenclature is:

Rational Evidence of Validity
- Face Validity
- Content Validity
- Factor Validity

Empirical Evidence of Validity
- Congruent Validity
- Concurrent Validity
- Predictive Validity
- Construct Validity (Shertzer & Linden, 1979)

This classification closely resembles the one devised by Technical Recommendations 25 years ago. There are, however, three notable

additions: face validity, factor validity and congruent validity.

Face validity assesses the degree to which the test seems relevant to the subject taking it, rather than assessing the validity of the test. Factor validity is an adjunct to content validity. With the advent of factor analysis, test constructors can now determine how closely their test follows the test plan. Congruent validity measures the degree of correlation between the test being studied and others designed to assess the same trait.

The remaining four have remained relatively unchanged over the past 25 years. Their definitions are virtually the same as those outlined by Chronbach and Meehl (1955). In this study, Chronbach and Meehl's (1955) definition of

Predictive Validity

Concurrent Validity

Content Validity

Construct Validity

will be used, while Shertzer and Linden's (1979) definitions for

Face Validity

Factor Validity

Congruent Validity

will be considered.

In light of the foregoing, it can be seen that the concept of validity has been in existence for many years and that to be of any value, a psychological test must demonstrate validity.

Validity of the Aptitude Inventory

The Aptitude Inventory is a personality test designed to measure the following personality variables; Intelligent Job Performance (I),

Leadership Qualities (L), Proper Job Attitude (A), Relations with Others (R). For more detail, see page 16.

Can the Aptitude Inventory demonstrate validity? Even a decade after the test was published, the revised Manual (Denton, 1969) lists extremely limited validity research. In this manual, three published and four unpublished studies are cited. In a personal communication, Denton (1979) stated "no further published research is available." He also said "I have instructed Buros to publish no further reviews of my tests."

When referring to content validity, Denton (1969) states "there can be little doubt of the content validity of the A.I." (p. 10). This statement is based on the fact that the items were selected from a "universe of ratings" used by management personnel, and the phrases were lifted word for word, changed only to the 1st person singular. This is, at best, questionable evidence of content validity. To measure construct validity, the test author correlated the A.I. with two instruments also published by his company, Psychological Business Research, the "P.B.R. Personal Survey" and "Biographical Information Blank." Denton (1969) summarizes the findings of these studies by stating "the correlations in this matrix are of expected magnitude and are in the proper direction" (p. 11). He does not, however, comment on the very low correlations (.02 to .33).

In a published study (Ross, 1963), construct validity as well as concurrent validity of the A.I. were to be determined. Scores on the A.I. were correlated with five factors obtained by factor analyzing a graphic rating form. The criterion factors determined were:

1. Over-all Sales Ability

2. Control of Assignment Details
3. Managerial Potential
4. Motivation
5. Capability in Systems Studies (p. 285).

The following correlation matrix was obtained:

	I	L	A	R	1	2	3	4	5	N=96
I	.01	-.06	-.11		.15	.30*	.32*	.-03	.08	
L		.24	-.21		.08	.-11	.23	.02	.05	
A			-.28		.12	.-10	.04	.27*	.14	
R					.05	.-03	.-03	.-01	.-06	(p. 287)

Based on this research, Denton in Ross (1963) concludes "Enough data are available for the A.I. that one would not begin with the hypothesis that the true population validities are zero" (p. 287). This is a rather strong statement, when there are only three correlations significant at the $p \leq .01$ confidence level.

In the A.I. manual, Denton (1969), reports a study by Guion (1963) and concludes "The results of this study ... seem most indicative of the construct validity of the A.I. ..." (p. 12). Guion (1963), however, concludes that because of the small sample size (48), the results are not significant. Denton's (1969) concluding comment in dealing with construct validity is "from the four studies reported, it seems apparent that the A.I. has construct validity." A more logical conclusion appears to be that more research is required to determine the construct validity of the A.I.

To demonstrate concurrent validity, 39 junior salesmen and 69 senior salesmen were rated by their superiors on the four A.I. scales, and these ratings were correlated with A.I. results. Results for the

individual scales were .28, -.16, -.07, .15, with a multiple correlation of .35 for juniors, and .22, .19, .09, -.12 multiple correlation .31 for seniors. In both cases, only the multiple correlation is significant at the .01 level of confidence, suggesting that even someone who has a thorough knowledge of an individual's performance on the job cannot predict well his subordinate scores on the A.I. Two other unpublished concurrent validity studies are briefly reported in the manual. Correlations in these studies are:

I	L	A	R	
.10	.15	.38	-.15	
.17	.10	.21	-.03	(Denton, 1969).

The manual for the Aptitude Inventory states that the average validity for the A.I. scales are:

I	L	A	R
.141	.152	.290	.050

To arrive at these figures, the test author averaged the correlation coefficients that were available and came out with a validity coefficient. This is an extremely questionable practice that has no support in the literature. Further, the coefficients are very low, questioning the A.I.'s validity.

In the only predictive validity research, it was determined in an unpublished study of a large aircraft manufacturing organization that in a battery of aptitude achievement intelligence, interest and personality tests, the best predictor of success as a supervisor was the total score of the A.I.. Denton (1969) states that this study "provides solid evidence of the concurrent validity of the A.I. in the selection of supervisors" (p. 15). However, this study as well as all of the other reported

research on the A.I. demonstrates little proof of the validity of the individual scales.

From this overview of the validity information available to date, it can be seen that before it can be used with confidence, more conclusive evidence of A.I.'s validity must be obtained.

Personality Testing

To proceed with validation research of the A.I., it is helpful to consider a historical overview of psychological tests and to compare the A.I. with two other widely used instruments.

The origin of personality testing through self report measures can be traced back to the first World War. Robert S. Woodworth constructed the Woodworth Personal Data Sheet, the first of a series of adjustment inventories (Woodworth, 1951). Woodworth's inventory was designed to identify soldiers likely to break down in combat. Since psychiatric interviews of all recruits were not practical, Woodworth devised a list of symptoms which psychiatrists would check out in interview, from which he developed a paper and pencil questionnaire (Woodworth, 1951). This inventory "had appreciable power to detect maladjusted soldiers" (Chronbach, 1970, p. 521).

Woodworth's scale was followed by a number of "adjustment inventories that list problems and symptoms or grievances to be checked" (Shertzer & Linden, 1979, p. 321). These inventories are "screening instruments," and should be used to identify individuals requiring further assessment or therapy (Shertzer & Linden, 1979, Chronbach, 1970). The predominance of behaviouristic psychologists between 1920 and 1945 influenced test construction to the extent that the questions placed more emphasis on what an individual did than on

how he thought or felt (Shertzer & Linden, 1979). The personality questionnaires were "broadened to described separate aspects of behaviour in a trait profile" (Chronbach, 1970, p. 522).

The second generation of personality questionnaires appeared after 1940. These questionnaires attempted to examine a number of traits that make up an individual's personality. Shertzer & Linden (1979) define trait as "psychological realities that exist in some tangible form in the person" (p. 315 & 316). Professionals who accept trait theory "assume that particular traits are common to many people, that they vary in amounts and that they can be inferred by measuring behavioural indicators ...; (also that) traits are enduring predispositions that exert fairly generalized effects on behaviour" (Shertzer & Linden, 1979, p. 316).

For a variety of reasons, the construction of personality inventories has tapered off in recent years. However, should there be a forward movement in test construction, the third generation "must be a generation of construct development and validation" (Chronbach, 1970, p. 526).

The three personality inventories considered in this study are the California Psychological Inventory, The Edwards Personal Preference Inventory, and the Management Aptitude Inventory.

California Psychological Inventory

The C.P.I. is a personality test constructed by Harrison Gough for persons 13 years of age or older. It was developed to "provide a comprehensive, multidimensional personality description of normal persons in a variety of non-clinical settings" (Shertzer & Linden, 1979, p. 331). The C.P.I. provides standard scores on 18 scales, 3

of which are validity scales. Each scale "is designed to forecast what a person will say or do under defined conditions and to identify individuals who will be described in characteristic ways by others who know them well ..." (Gough, 1975, p. 5).

Gough groups the scale "for convenience" into four broad categories.

CLASS I. Measures of Poise, Ascendancy, Self-Assurance, and Interpersonal Adequacy

1. Do Dominance
2. Cs Capacity for Status
3. Sy Sociability
4. Sp Social Presence
5. Sa Self-acceptance
6. Wb Sense of Well-being

CLASS II. Measures of Socialization, Responsibility, Intrapersonal Values, and Character

7. Re Responsibility
8. So Socialization
9. Se Self-control
10. To Tolerance
11. Gi Good Impression
12. Cm Communality

CLASS III. Measures of Achievement Potential and Intellectual Efficiency

13. Ac Achievement via Conformance
14. Ai Achievement via Independence
15. Ie Intellectual Efficiency

CLASS IV. Measures of Intellectual and Interest Modes

16. Py Psychological-mindedness

17. Fx Flexibility

18. Fe Femininity (Gough, 1975, p. 5).

Although this grouping is similar to what has resulted from various factor analyses, the grouping was not attained factorily (Megargee, 1972).

Development of the scales was undertaken to maximize validity. Gough (1973) describes how "eleven of the 18 scales in the inventory were developed by utilizing empirically derived scoring weights assigned to the responses found to differentiate defined criterion groups ... four of the scores were originally judged by the author as indicating a designated variable and refined by internal consistency checks, ... and three scores are derived empirically to detect tendencies to fake" (p. 39).

Test retest reliability was measured on a sample of 200 prisoners retested after one to three weeks. Correlations range from .49 to .87, median .80. High school subjects retested after one year provide median test retest conditions of .65 for males and .68 for females.

The C.P.I. has been the subject of voluminous research. Buros (1978), cites the 1322rd reference on the instrument. Included in this list are numerous validity studies. Megargee's 1972 handbook reviews the 18 scales and the validity studies related to each, and concludes that the C.P.I. is the most validated test instrument available for use with a normal population.

Critical reviews in Buros (1959, 1965, 1972, 1978) by Kelly, Thorndike, Chronbach, Goldberg, Walsh, Crites, and Gyther point out

"convincing evidence to validate each of the 18 scales" (Lake et al., 1973, p. 39). They also point out problems in the development of the inventory: 1) High interrcorrelations among the scales; 2) Using only external groups in population samples; 3) Grouping scales for interpretive convenience, rather than by factor analysis.

There seems to be a general consensus in the literature that the C.P.I. is the best personality test for use with normals in the market today (Buros, 1959, 1965, 1972, 1979; Megargee, 1972; Shertzer & Linden, 1979).

Edwards Personal Preference Schedule

The E.P.P.S. was designed by Allen R. Edwards "to assess the relative importance to the individual of fifteen key needs or motives selected from H. A. Murray's need system (Shertzer & Linden, 1979, p. 324). The E.P.P.S. (1966) provides measures of the following 15 "personality" variables:

1. Achievement (ach)
2. Deference (def)
3. Order (ord)
4. Exhibition (exh)
5. Autonomy (aut)
6. Affiliation (aff)
7. Intraception (int)
8. Succorance (suc)
9. Dominance (dom)
10. Abasement (aba)
11. Nurturance (nur)
12. Change (chg)

13. Endurance (end)

14. Heterosexuality (het)

15. Aggression (agg) (Edwards, 1966, p. 5).

A forced choice method was adopted to inhibit falsification. In test construction, each one of the nine statements from each need area was paired with statements from the remaining need areas; each pairing was made twice, and 15 were repeated to provide a consistency score. Great care was taken to ensure that each pair was of equal social desirability (Edwards, 1953).

Internal consistency reliability ranged from .60 to .87 with a mean of .78 (Edwards, 1966). Test retest reliability with a one week interval, ranged from .74 to .87, median .78; with a three week interval, ranged from .55 to .87, median .73. Memory effects may have influenced the one week results (Lake et al., 1973).

Validity data in the Manual (Edwards, 1966) are sparse. Edwards (1966) describes studies where individuals rated themselves on the 15 needs; the researcher then compared the obtained scores with E.P.P.S. results. Edwards also describes correlation studies with instruments such as the Guilford Martin Personnel Inventory, and the Taylor Manifest Anxiety Scale. There is, however, little evidence of validity in any of the published reviews.

The "E.P.P.S. has received considerable criticism for its insufficient validation to which Edwards, even on his revised manual (1966) has not yet replied" (Lake, et al., 1973, p. 70). Critics of the Edwards (Buros, 1959, 1965, 1972; Lake, et al., 1973; Shertzer & Lindon, 1979) emphasize two points:

1. The scores obtained are ipsative. They reflect the relative

strengths of the needs within the individual, rather than the strengths of these needs relative to the needs of other individuals.

2. Even though the instrument has been greatly criticized for lack of validity research, the revised manual (1966), does not provide any more validity research than its predecessor (1959).

Majority opinion in the literature appears to be toward recommending the E.P.P.S. for research, but questioning its utility as a clinical instrument.

Aptitude Inventory

The A.I. is published in the same form under three titles, Management Aptitude Inventory, Employment Aptitude Inventory, Sales Aptitude Inventory (Denton, 1969). Each test has identical items, and identical norms. The A.I. provides measures of the following psychological personality variables.

1. Intellectual Job Preference (I)
2. Leadership Qualities (L)
3. Proper Job Attitude (A)
4. Relations with Others (R).

Items in the A.I. were derived from merit rating studies conducted by Denton. They have been worded in the first person singular, but otherwise are almost verbatim descriptions of good and poor performance as described by senior management in several companies. Seventy-five of the items are of a forced choice nature, while twenty-three are yes/no questions.

Reliability data comes from background research studies in a

prepublication model of the inventory. Test retest reliabilities range from .61 to .66. This data is of little value in evaluating the present test form which was changed from the original model by increasing the number of items, editing the questions, and revising the scoring keys. Jurgensen, in Buros (1965), comments on these test retest reliabilities of 113 college students over a two week period; "Many test experts will not agree with the author that reliabilities of this magnitude are acceptable" (p. 1249).

As stated previously, even in the revised manual (1969), there is little validity research.

Critics of the A.I. are concerned with its limited reliability and validity information (Buros, 1969). It has not been recommended for any situations but those where extensive research is possible prior to its implementation (Siegel, 1959).

Similarities of these Instruments

Though the Aptitude Inventory, the California Psychological Inventory and the Edwards Personal Preference Schedule have different titles for their respective scales, from the definitions listed in their manuals, they appear to measure conceptually similar dimensions. These apparent similarities will be given a three part discussion:

1. the opinions of two "experts" regarding which scales measure similar traits;
2. the definitions of the scales from their manuals;
3. the related literature.

Two psychologists who use these three instruments on a daily basis outlined the scales they felt measured similar traits. The two "expert" psychologists are associates of a firm of Industrial

Psychologists. One holds psychological registration in three Canadian provinces, has over 15 years experience in executive appraisal, and many more years clinical experience. The second has a B.A. (Honours) degree, with over 10 years experience in executive appraisal, the first two years under supervision of Clinician one in what can be considered an intensive internship.

The scales which these "experts" felt measured similar traits were as follows:

	Attitude Inventory	Edwards Personal Preference Schedule	California Psychological Inventory
(I)	Intelligent Job Performance	Achievement	Intellectual Efficiency
			Achievement via Independence
(L)	Leadership Qualities	Dominance	Dominance
		Autonomy	Capacity for Status
		Exhibition	Social Presence
			Self acceptance
(A)	Proper Job Attitude	Endurance	Sense of Well-being
			Responsibility
(R)	Relations with Others	Affiliation	Sociability
		Nurturance	Socialization
			Self control
			Tolerance

Table 1 presents the definitions for each scale of the three instruments of interest to this study, presenting a theoretical rationale for the experts' expectations. By viewing Table 1, it is evident that the scales outlined above appear to measure very similar dimensions.

Table 1

Brief descriptions, sample phases and manifest needs of the four A.I. scales and conceptually similar scale of the E.P.P.S. and C.P.I.

Attitude Inventory	Edwards Personal Preference Schedule	California Psychological Inventory
Intelligent Job Performance - eager to learn, thorough on job assignments, alert, open-minded, intellectually curious	Achievement - to do one's best to accomplish tasks requiring skill and effort, to do a job well, to solve difficult problems	Intellectual Efficiency - Intellectually efficient, clear thinking, alert, placing high value on intellectual matters
		Achievement via Independence - Achievement via autonomy and independence, mature, forceful
Leadership Qualities - lead a group, able to gain respect, able to motivate associates, organize and delegate	Dominance - to be a leader of groups, to be regarded by others as a leader	Dominance - leadership ability, dominance, persistence
	Autonomy - to be independent, to feel free to do what one wants	Capacity for Status - capacity for status, ambitious, active, forceful
	Exhibition - to say witty and clever things, to be the centre of attention	Social Presence - poised, spontaneous, self-confident
		Self Acceptance - sense of personal worth, self-assured, outspoken
Proper Job Attitude - ambitious and hard working, willing to work hard, drive, ambition, determination	Endurance - to keep at a job until it is finished, to work hard at a task, to put in long hours of work	Sense of Well-Being - energetic, alert, ambitious, valuing work for its own sake

Table 1 Continued

Attitude Inventory	Edwards Personal Preference Schedule	California Psychological Inventory
		Responsibility - conscientious, responsible, planful and thorough
Relations with Others - get along well with others, tolerant, not critical, fair and impartial	Affiliation - to be loyal to friends, to participate in friendly groups, to form strong attachments	Sociability - outgoing, sociable, participative
	Nurturance - to be with others when they are in trouble, to be generous with others	Socialization - socially mature, serious, honest, conforming
		Self Control - calm, patient, honest, and conscientious
		Tolerance - tolerant, accepting, non-judgmental, clear thinking

Though research articles for both the E.P.P.S. and C.P.I. number in the thousands, there are very few studies that correlate the two tests. McKee and Turner (1961), correlate scores on the E.P.P.S. and the C.P.I. with ratings of "drives" made fifteen years earlier by judges who knew the subjects "extremely well." In this article, McKee et al (1961) report correlations for ten of the twenty-four scales the "experts" felt would correlate moderately. These correlations however, are between judges' ratings on the needs and C.P.I. scores. Of the ten correlations reported, six for the male sample, and seven for the female sample were significant in the expected direction.

Dunnette, Kirchner, and DeGedio (1958), correlated the C.P.I. with the E.P.P.S. and each in turn with occupational area scores obtained from the Strong Vocational Interest Blank. Their sample consisted of 102 employers of the Minnesota Mining and Manufacturing Company, including thirty sales managers, thirty-eight salesmen, nineteen project supervisors and fifteen project engineers. In this study, Dunnette et al present only significant ($p \leq .05$ level) correlations; they found 11 of the 24 correlations were positive and in the expected direction. Table 2 outlines these correlations.

Inspection of Table 2 shows that the significant correlations for the E.P.P.S. and C.P.I. range from .20 to .41. Dunnette et al (1958) state that "a brief examination of these scales shows that the direction of the association among the various variables makes good 'clinical sense'" (p. 179). Should the A.I. demonstrate construct validity, the interrelations between itself and the C.P.I. and E.P.P.S. should also

Table 2

Significant Correlations Between Scales of E.P.P.S. and of C.P.I. (Dunnette, et al, 1958)

C.P.I. Scales	Achievement	Dominance Autonomy Exhibition	Endurance	Affiliation Nurturance
Intellectual Efficiency	X			
Achievements via Independence	X			
Dominance		.34 x .34		
Capacity for Status		.26 x .23		
Social Presence		.41 x .35		
Self acceptance		.29 x .29		
Sense of Well-being			.23	
Responsibility			.23	
Sociability				X X
Socialization				X X
Self Control				X X
Tolerance				X .20

X indicates nonsignificant correlation

make "good clinical sense." That is, the intercorrelation should be in the direction predicted by "expert psychologists," and in the range obtained by Dunnette et al (1958).

Problems with Personality Tests

The three inventories considered in this study are all self report in format. The following issues have been raised regarding accepting this type of questionnaire as valid self-descriptions.

Often, items included in self report inventories are ambiguous. Chronbach (1970) and Shertzer & Linden (1979) cite the sample question "Do you make friends easily?" to demonstrate the ambiguity of words such as friends, and easily. Further, many self report inventories ask if a phrase applied to the individual always, frequently, or rarely. Simpson (1944), in a quantitative analysis of such words asked students what they meant by saying they did something usually. There was a large range in their responses, 25% of them replied that it meant over 90% of the time, while another 25% said under 70%. In discussing this research, Chronbach (1970) states "It is evident that subjects with identical behaviour choose very different adverbs to describe what they do" (p. 794).

A second problem occurs as many personality questionnaires can be falsified by the individual being tested. In a clinical setting, the subject may want to avoid a threatening diagnosis (Shertzer & Linden, 1979), whereas in a business setting, often the candidate's first concern is to obtain employment (Chronbach, 1970). Studies have demonstrated that personality questionnaires can be faked (Longstaff, 1948; Wesman, 1952; Dunnette, 1962); a number of studies have demonstrated that some questionnaires can be faked in a desired direction

(Longstaff, 1948).

To control the tendency to fake, two main approaches have been undertaken. Questionnaires such as the M.M.P.I. and the C.P.I. have included in them various "lie scales" to determine if an individual is attempting to fake good, or fake bad. The second approach is the forced choice questionnaire such as the E.P.P.S. and the A.I.. Those in favour of the forced choice technique suggest that this technique does not eliminate the risk of falsification, rather it reduces the ability of the test taken to produce a desired result (Shertzer & Linden, 1979). Critics, on the other hand, state that with forced choice questionnaires:

- results are invalid as often "the individual may have no logical basis for making a choice" (Shertzer & Linden, 1979, p. 346);
- the obtained scores cannot be manipulated by conventional arithmetic procedure, therefore norms should not be provided (Bauernfeind, 1962);
- at best, the ratings are ipsative, that is, "each person's scores are distributed around his own average, and therefore the analysis of his scores or ratings is meaningful only when interpreted in terms of his own performance ..." (Kleinmuntz, 1967, p. 206-207).

The force choice method does control for faking, as it ensures the individual cannot produce a desired result. However it greatly affects the interpretive value of the inventory.

A third problem with self report questionnaires is acquiescence. Some people have a generalized tendency to be agreeable; this characteristic can greatly affect true/false questionnaires. To overcome this

problem, many test developers attempt to key as many true responses as false responses in any given scale.

From the foregoing, we see that there are many problems with personality measures, even when they are well constructed and well validated. These problems, however, are compounded when an inventory does not have demonstrated validity. As mentioned previously, to ensure validity, a basic requirement is predictive validity.

Prediction

A required characteristic of a good psychological test is the ability to predict a criterion. Therefore, to be worthwhile, a test must demonstrate the ability to predict measurable results on the criterion. The "prediction problem" has been seen by Gough (1962) as the basic problem of applied psychology. There are two methods of prediction based on test results, clinical judgement and actuarial prediction.

Models of Clinical Judgement

The three general models for clinical decision making are the proper linear model, the improper linear model and the non-linear model (Goldberg, 1971; Dawes, 1979). In the proper linear model, the weights given to predictor variables are statistically chosen to optimize the relationship between the predictor and the criterion. Such models include standard regression analysis and discriminant function analysis. In the improper linear model, weights given to predictor variables are obtained by some non optimal way, such as by intuition or clinician prediction, or even by random selection. Non-linear models usually involve some type of moderator variable effect, such as when the weighting of one variable deviates according to the magnitude of the relationship between two or more other variables (Goldberg, 1971).

The Validity of Clinical Judgement

Research in the area of clinical judgement has focused most attention on the predictive validity of clinical decision making, specifically the difference between clinical and actuarial predictions. Various models of clinical judgement or prediction have been outlined above. Actuarial prediction, on the other hand, occurs when there are explicit rules by which predictions are made about individuals on the basis of experimental or statistically demonstrated associations between specified data and the criterion being predicted (Marks & Seeman, 1963).

Paul Meehl's (1954) now classic book, Clinical versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence, summarized studies dealing with the validity of clinical and actuarial decisions based on prediction of numerical criterion variables from numerical predictor variables. Meehl (1954) summarizes his findings by stating that of the "20 studies involving a comparison of clinical and actuarial methods, in all but one, ... the predictions made actuarially were either approximately equal or superior to those made by the clinician" (p. 119).

Virtually all of the literature since Meehl's (1954) book supports his above generalization. Wiggins (1973), in discussing the 50 studies reviewed by Meehl (1954, 1957 and 1965), states that "33 of them demonstrated the superiority of statistical over clinical data combinations, and the remaining 17 studies indicated that the two methods were approximately equal in predictive accuracy" (p. 221).

Holt (1958, 1970), criticized details of several of these studies, and sees the dichotomy of clinicians versus actuaries as artificial.

Dawes (1979), agrees "People are important ... it is always the individual who chooses variables, ... it is the human judge who knows the directional relationship between the predictor variables and the criterion of interest" (p. 573). Sawyer (1966) also emphasizes that the clinician is important, that his/her real strength is in bringing additional non psychometric information to the decision making process. Critics feel that these studies have compared actuarial prediction at its best with clinical prediction that is not at its best (Holt, 1958, 1970; Sawyer, 1966).

Dawes (1979) summarized this issue of clinical versus actuarial prediction by stating:

> Proper linear models work for a very simple reason. People are good at picking out the right predictor variables and at coding them in such a way that they have a conditionally monotone relationship with the criterion. People are bad at integrating information from diverse and incomparable sources. Proper linear models are good at such integration when the predictors have a conditionally monotone relationship to the criterion. (p. 574)

A surprising finding is that non linear models have also been demonstrated to be better predictors than clinical judgement. One method, "bootstrapping" (Dawes & Corrigan, 1974; Goldberg, 1970; Dawes, 1979) involves the building of a proper linear model based on a clinician's judgement about an outcome criterion and using that model in place of the judge. Dawes (1979) reviewed five studies involving a clinician, bootstrapping, random model, and an optimum model. The final three models consistently outperformed the clinical judgement.

Nonetheless, Scissons (1976) states that "the real problems of the predictive validity of clinical or actuarial judgement may be escaping both the clinician and actuary" (p. 12). Ash & Kroeker (1975), considering the efficacy of both models, indicate that a criterion predictor of .60, which is higher than most clinical or actuarial studies, is still appallingly low.

Consistent throughout the literature is the finding that actuarial prediction in the vast majority of cases outperforms clinical prediction. Therefore, in validity studies one must be aware of this observation and use both clinical and statistical methods. Regarding the problem of "appallingly low" correlations in personnel selection and/or vocational counselling, even with correlations of .6, accounting for 36% of the variance is far better than accounting for 0% of the variance. In addition, for the majority of studies regarding prediction, a criterion is often extremely difficult to measure; e.g., for the criterion job success, one can use supervisor ratings that may not be consistent, or direct productivity which may be affected by a number of variables. Therefore, the criterion may not be reflecting totally what it is expected to reflect.

Predicting Validity Studies of the A.I., E.P.P.S. and C.P.I.

Though the A.I. is also known as the Sales Aptitude Inventory, only one validity study has dealt with this test and a sales population. This study (Denton, 1969), however was not a prediction study, rather it was a study of concurrent validity.

Although no research is available regarding the ability of the E.P.P.S. or C.P.I. to predict success in sales, some research has been conducted regarding predicting job success. "The E.P.P.S. has been

used successfully for predictive purposes in several studies" (Slocum & Hand, 1971, p. 28). Norrell & Grater (1960) demonstrate that prediction of vocational interests is possible using the E.P.P.S.; Broomer (1962) found that three E.P.P.S. scales were significantly related to foremen's productivity, and Kuhlen (1963) reported significant correlations between E.P.P.S. needs and occupational satisfaction for teachers. Slocum et al (1971) attempted to predict job performance and employee satisfaction from the E.P.P.S.. "Zero order correlation between E.P.P.S. scores, the five measures of employee satisfaction (confidential questionnaire) and the eight measures of job performance (supervisor ratings) were generally small and most were statistically insignificant" (Slocum, et al, 1971, p. 30). Correlations ranged from .-34 to .33. Multiple regression, and corrected R^2 values were also small, the latter ranging between .057 and .305.

Muchinsky and Hoyt (1973) on the other hand, did not find need for achievement related to vocational accomplishment and did not find affiliation related to people-oriented accomplishments in predicting vocational performance of engineers. They conclude "a need and its behavioural counterpart should not be assumed to be related in all situations" (Muchinsky & Hoyt, 1973, p. 123).

Kelly (1974) used the C.P.I., the E.P.P.S., the 16 P.F., and the M.M.P.I. to predict leadership in nursing. Multiple regression scores to predict promoted or not promoted range from .42 to .68. Kelly further uses a double cross validation and only one test, the 16 P.F., significantly ($p \leq .05$) predicts the criterion.

Individual scales of the C.P.I. as well as the whole instrument have also been used successfully for predictive purposes. Gottfredson

& Lipstein (1975) found that the C.P.I. "So" scale adds significantly to the prediction of Parolee and Probationer Employment stability. Dorin (1974) found that the C.P.I. was a valuable addition to the selection of residence hall staff. Preston, in a 1976 study, found the C.P.I. to successfully predict police performance. Using 17 of the 18 C.P.I. scales, he obtained a multiple regression coefficient of .61, R^2 of .37 with the highest simple correlation being with the well being (WB) scale, .40.

Summary and Research Questions

It is evident that validity is a key concept in test construction. All too often, a psychological instrument is used widely when very little evidence supports its validity. The Aptitude Inventory is one such instrument. There is a definite need to add to the validation research on this test.

The questions posed by this study are asked to add to the research on the construct and predictive validity of the Aptitude Inventory.

Question 1

Are relationships among scales of the A.I., E.P.P.S. and C.P.I. in the direction and of the magnitude expected by two "experts" and supported by the literature?

Question 2

To what extent can a clinician predict a candidate's "success" in real estate using the A.I.?

Question 3

To what extent can a clinician predict a candidate's "success" in real estate sales using ten psychological tests and interview notes?

Question 4

To what extent can scales of the A.I. predict a candidate's "success" in real estate sales?

Question 5

To what extent can scales of the C.P.I. predict a candidate's "success" in real estate sales?

Question 6

To what extent can sales of the A.I., E.P.P.S. and C.P.I. together predict "success" in real estate sales?

Question 7

Which methods described in Questions 2 through 6 best predict sales volume in real estate sales?

Question 8

Does the A.I. add to the actuarial prediction of "success" in real estate sales in a multiple regression when results are available for the A.I., E.P.P.S. and C.P.I.?

Question 1 investigates the construct validity of the A.I. while Questions 2 through 8 investigates the predictive validity of the Aptitude Inventory, the Edwards Personal Preference Schedule and the California Psychological Inventory.

CHAPTER 3

PROCEDURE AND DESIGN

In order to pursue the questions posed in Chapter two, this investigation was broken into two components, Study A and Study B. In this chapter, the data gathering procedure, sample and analysis procedure for each study are presented and discussed. The final section is a discussion of the limitations of these studies.

Study A: Construct Validity

The A.I., E.P.P.S., and C.P.I. are three personality tests that have some overlapping scales, or scales that purport to measure similar dimensions. Moderate correlations (.4 to .6) between scales on the A.I. and conceptually similar scales on the E.P.P.S. and C.P.I. would be one indicator of the construct validity of the poorly researched A.I..

Data Gathering Procedure

One active researcher in Canada collected demographic information and psychological test results from 3 firms of industrial psychologists located in Edmonton, Calgary, Saskatoon, and Toronto. The data collected by this researcher includes the correlated scores on a number of psychological tests, including the three personality tests considered in this study. These data were made available for this study.

Sample

The sample consists of 2425 recruitment and appraisal candidates tested by the participating psychologists across Canada. Recruitment candidates are individuals who have applied to senior positions through the recruiting division of industrial psychologists. Appraisal candidates are individuals sent to the psychologists by their own employers or prospective employers for assessment in order to

determine strengths, weaknesses and future development potential. The sample was 100% male, ranging in age from 18 to 63 with a mean age of 35.18. The subjects were from a variety of occupations including Engineers, Accountants, Salesmen, Technicians, Personnel Technicians, Marketing and Senior Management Personnel. Regarding level of applicant's occupations, 38.4% were junior, 57.8% senior, and 3.8% Vice-President and up.

Analysis Procedure

The results of the A.I., E.P.P.S., and C.P.I. were analyzed using Pearson Product Moment Correlation, and Factor Analysis.

Study B: Predictive Validity

Since the A.I. is also known as the Sales Aptitude Inventory it is purported to be able to predict success as a sales person. This study will determine the predictive validities of the A.I. and two other well known tests, the C.P.I. and E.P.P.S., and will determine if the A.I. is a useful tool in predicting success in real estate sales.

Data Gathering Procedure

Data was gathered on 22 real estate salesmen who completed their first year of sales. An Edmonton real estate company refers prospective hires to A. W. Fraser & Associates for appraisal to determine their potential for success in this field. The candidates are interviewed by a consultant and administered a battery of psychological tests. The appraisal includes 10 psychological tests, and takes approximately 6 hours. Selected candidates are hired by the real estate company, and production records of total dollars brought in by the sales person are recorded monthly. The criterion for success

in real estate sales in this study is based on production records (adjusted for inflation) for the first twelve months as a sales person.

The following data were obtained on all candidates:

1. Interview
2. C.P.I. results
3. E.P.P.S. results
4. A.I. results
5. Remaining 7 psychological tests
6. Production records.

Items 1 through 5 were collected at time of initial appraisal. Item 6 was obtained later from the participating real estate company.

As the date of commencement for the sales persons ranged from October, 1977 to December, 1978 (Table 3), the production records were adjusted for inflation. To determine accurately the average monthly increase for homes in Edmonton for the time period of this study, a second Edmonton real estate firm undertook a market survey. The results of this survey (Table 4) indicate that the average monthly increase for homes in a typical Edmonton neighbourhood was .7% for the time period September, 1977 to October, 1979. Table 3 shows the commencement date, production records and production records adjusted for inflation (criterion) for all 22 subjects.

After the test results and production records (adjusted for inflation) were collected, a clinician made 2 predictions for success in real estate based on:

1. the A.I. alone;
2. 10 test battery plus interview.

Table 3

Production Records for Twenty-Two
Real Estate Salespersons

Month Started	Month Ended	Production Records	Monthly Adjustment & Percent Addition	Production Record Adjusted for Inflation
1 Aug. 78	July 79	$ 78,699.30	4 (.28%)	$ 78,999.66
2 May 78	Apr. 79	61,844.68	6 (.42%)	62,104.43
3 Jan. 78	Dec. 78	133,778.58	11 (.77%)	134,808.67
4 May 78	Apr. 79	18,967.00	6 (.42%)	19,020.11
5 Oct. 77	Sept. 78	52,051.05	14 (.98%)	52,561.15
6 Dec. 77	Nov. 78	90,739.40	12 (.84%)	91,501.61
7 Oct. 78	Sept. 79	49,387.20	2 (.14%)	49,456.34
8 Jan. 78	Dec. 78	37,085.89	11 (.77%)	37,370.55
9 Oct. 78	Sept. 78	34,784.00	2 (.14%)	34,832.70
10 Mar. 78	Jan. 79	20,580.00	9 (.56%)	20,695.25
11 Aug. 78	July 79	36,678.75	4 (.28%)	36,781.45
12 Mar. 78	Feb. 79	28,971.96	9 (.56%)	29,133.24
13 May 78	March 79	18,975.65	7 (.28%)	19,028.78
14 Jan. 78	Dec. 78	55,003.50	11 (.77%)	55,427.03
15 Nov. 78	Oct. 78	29,689.05	1 (.07%)	29,709.05
16 Dec. 77	Nov. 78	41,551.98	12 (.84%)	41,901.02
17 Dec. 78	Oct. 79	22,484.75	0 (0)	22,484.75
18 Sept. 77	Aug. 78	47,333.60	15 (1.05%)	47,830.60
19 Aug. 78	July 79	36,600.49	4 (.28%)	36,702.97
20 Dec. 78	Oct. 79	20,196.00	0 (0)	20,196.00
21 Sept. 78	Jan. 79	52,904.30	10 (.70%)	53,274.63
22 Oct. 78	Sept. 79	26,144.12	2 (.14%)	26,180.72

Table 4

Market Survey of Real Estate Prices
from September, 1977 to October, 1979

Date	Address	Size (Sq. Ft.)	Sale Price	Cost Per Sq. Ft.	
Sept. 77	6704 - 94A Ave.	1100	$72,000	65.5	(65.5)**
Oct.					
Nov.	8407 - 56 St.	1120	73,500	65.6	(66.4)
Dec.	9620 Ottewell Rd.	1138	77,000	67.7	(66.8)
Jan. 78	6304 - 94B Ave.	1098	74,500	67.8	(67.3)
Feb.	6303 - 94B Ave.	1185	81,000	68.4	(67.8)
Mar.	5312 - 89 Ave.	1070	73,500	68.7	(68.3)
Apr.	8821 - 74 St.	1120	76,000	67.9	(68.7)
May	5303 - 94B Ave.	1088	75,300	69.2	(69.2)
June	9523 - 52 St.	1095	77,500	70.8	(69.7)
July	7904 - 91 Ave.	1070	73,000	*68.3	(70.2)
Aug.	6112 - 94A Ave.	1063	74,000	69.6	(70.7)
Sept.	9103 - 72 St.	1175	81,500	69.4	(71.2)
Oct.					
Nov.					
Dec.	9832 - 76 St.	1184	84,000	*70.9	(72.7)
Jan. 79	9611 - 52 St.	1160	85,500	73.7	(73.2)
Feb.	9831 - 68 St.	1187	86,000	*72.4	(73.7)
Mar.	5508 - 97 Ave.	1175	86,000	73.2	(74.2)
Apr.	5507 - 90A Ave.	1100	80,500	73.2	(74.7)
May	9903 - 76 St.	1070	79,900	74.7	(75.3)
June	9527 - 85 St.	1150	87,000	75.7	(75.8)
July	7404 - 94B Ave.	1140	87,500	76.7	(76.3)
Aug.	9522 - 83 St.	1040	80,500	77.4	(76.8)
Sept.	9112 - 79 St.	1093	82,500	*75.5	(77.4)
Oct.	9624 - 52 St.	1050	81,900	78.0	(77.9)

* Single garage only, i.e. if typical with double garage sale price would be about $1750 higher and cost per sq. ft. would be about $1.50 greater.

** Using $65.5/sq. ft. as the cost of the typical home in Sept. of 1977 the figures shown bracketed represent a per month increase of 0.7%.

Table 4 Continued

The above information is based on actual records of the sales of properties so described, i.e. Ottewell/Holyrood located 3 bedroom bungalows of 1120 sq. ft. average size with developed basements, above average locations and lots, double detached garages and no fireplace.

The clinician involved holds psychological registration in three Canadian provinces, has over 15 years experience in executive appraisal, and many more years clinical experience.

The participating clinician was provided the anonymous results of the A.I. (Table 5) and asked to rate the candidates on the following 5 point scale:

1.	Very good	75% chance of success	25% chance of failure
2.	Good	60% chance of success	40% chance of failure
3.	Average	50% chance of success	50% chance of failure
4.	Fair	40% chance of success	60% chance of failure
5.	Poor	25% chance of success	75% chance of failure

Numbers 1 through 22 in Table 5 refer to the same candidates as 1-22 in Table 3. After the clinician completed this rating, he was provided the results of all the tests in the sales appraisal and the notes of an interview with each subject. The clinician was asked to rate the subjects on the same 5 point scale.

Sample

The sample consists of 22 real estate sales persons who were appraised before embarking on a career in real estate sales. The sample included 16 men and 6 women, ranging in age from 24 to 48 years with a mean age of 34.47. Ninety percent of the sample was married with a mean number of 1.63 children; 5 percent were separated and 5 percent were single.

Analysis Procedure

1. Pearson Product Moment Correlations were calculated for clinician's prediction based on the AI and the criterion (success in real estate sales), as well as for the

Table 5

Aptitude Inventory Results for Twenty-Two Real Estate Salespersons

	Intelligent Job Performance	Leadership Qualities	Proper Job Attitude	Relations with Others	Rated 1-5
1	90	80	60	40	
2	20	01	05	20	
3	80	70	80	70	
4	30	01	80	95	
5	80	90	60	10	
6	50	60	70	80	
7	05	60	60	30	
8	30	70	80	90	
9	10	30	95	30	
10	60	70	30	80	
11	70	30	90	50	
12	70	30	20	60	
13	20	05	20	80	
14	60	30	20	90	
15	01	90	05	80	
16	40	05	70	10	
17	80	01	90	40	
18	05	10	01	60	
19	40	90	10	95	
20	40	05	05	70	
21	01	80	90	30	
22	10	01	05	90	

clinician's prediction based on the entire sales appraisal battery and the criterion.

2. Multiple regressions were calculated between the criterion and
   a) the 4 AI scales
   b) the 15 EPPS scales
   c) the 13 CPI scales.
3. A multiple regression analysis was completed for the 3 best predictors for each of the AI, EPPS and CPI and the criterion.

In procedures 2 and 3, a Step wise multiple regression analysis was undertaken. Multiple Regression and R^2 values were given step wise for the single best predictor, and for all other combinations of predictors up to and including all predictors. The number of scales to be considered in the multiple regression was determined by using an Analysis of Variance to determine the scales that significantly added to the Multiple Regression at the .05 level. To get the appropriate F value, the following formula is used.

$$F = \frac{(R_f^2 - R_r^2) / (m_1 - m_2)}{(1 - R_f^2) / (N - m_1)}$$ where:

$R_f^2 = R^2$ of the additional predictor

$R_r^2 - R^2$ of the original predictor(s)

m_1 = number of variables for R_f^2

m_2 = number of variables for R_r^2

N = sample size.

Limitations

1. Subject selection for Study A was not random. However, subjects can be considered to be representative of clients

who undertake appraisal by industrial psychologists.

2. The clinician and all subjects in Study A are male. This precludes generalization of results to female populations.
3. No information will be forthcoming regarding the predictive validity of AI beyond predicting success in real estate sales.
4. Subject selection for Study B was not random.
5. The Study B sample of 22 does not allow for the preferred method of analysis, double cross validation. Further, with a sample of this size, the regression equation obtained for predicting success in real estate sales will have a large standard error. Further research is required to obtain a valid regression equation.

CHAPTER 4

FINDINGS OF THE STUDY

Chapter 4 reports the findings of studies A and B. Each question posed relating to first construct validity, then to predictive validity is reported, followed by a discussion of the findings.

Study A: Construct Validity

Question 1

Are relationships among scales of the Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory in the direction and of the magnitude expected by two "experts" and supported by the literature?

Table 6 is a summary of the correlations that are expected to be positive and moderate (see Appendix 1 for all of the intercorrelations among the A.I., E.P.P.S. and C.P.I.). By examining Table 8 it is evident that the correlations between A.I. scales, and conceptually similar scales of the E.P.P.S. and C.P.I. range between -.07 and .32. Appendix 1 indicates that the 132 correlations between the four A.I. scales and all scales of the E.P.P.S. and C.P.I., range between -.23 and .32 with 83, or 63% being between -.1 and +.1.

Table 6

Correlations Between Conceptually Similar Scales of the A.I., the E.P.P.S. and the C.P.I.

E.P.P.S. and C.P.I. Scales	Aptitude Inventory Scales			
	Intelligent Job Performance	Leadership Qualities	Proper Job Attitude	Relations with Others
Achievement	.17			
Dominance		.32		
Autonomy		.06		
Exhibition		.13		
Endurance			.21	
Affiliation				.26
Nurturance				.27
Intellectual Efficiency	.11			
Achievement in Independence	.17			
Dominance		.29		
Capacity for Status		.20		
Social Presence		.18		
Self acceptance		.19		
Sense of Well-Being			-.01	
Responsibility			-.07	
Sociability				-.03
Socialization				.03
Self control				.03
Tolerance				.07
	n=2425			

Discussion

The above correlations are very low, indicating that the A.I. does not measure the same dimensions as do the E.P.P.S. and the C.P.I..

Additional Findings Pertinent to Question 1

To determine if there is any subtle relationship among scales of the A.I., E.P.P.S. and C.P.I., factor analysis (with varimax rotation) was completed. By inspection of Table 7 it is apparent that the scales that load highest (.4 or greater) on the individual factors are:

Factor 1	Factor 2	Factor 3	Factor 4
C.P.I. Wb	C.P.I. Do	C.P.I. To	E.P.P.S. Aff
C.P.I. Re	C.P.I. Cs	C.P.I. Ai	E.P.P.S. Nur
C.P.I. So	C.P.I. Sy	C.P.I. Ie	
C.P.I. Sc	C.P.I. Sp	C.P.I. Fx	
C.P.I. To	C.P.I. Sa		
C.P.I. Gi	E.P.P.S. Dom		
C.P.I. Ac			
C.P.I. Ie			

These data indicate that there are no A.I. scales, and only 1 E.P.P.S. scale loading significantly on the first three factors which are predominated by C.P.I. scales.

Discussion

Again, it appears that the A.I. scales do not measure the same dimensions as the E.P.P.S. and the C.P.I..

Table 7

Varimax Rotated Factor Matrix of the
A.I., E.P.P.S. and C.P.I.

Scale	Factor 1	2	3	4
Intelligent Job Performance	.268	.030	.048	.041
Leadership Qualities	.102	.322	-.128	-.129
Proper Job Attitude	-.077	.083	-.154	-.082
Relations with Others	.022	-.023	.054	.251
Dominance	.177	.735	.007	-.019
Capacity for Status	.252	.585	.266	.052
Sociability	.217	.789	.026	.059
Social Presence	-.082	.680	.340	.003
Self acceptance	-.202	.708	-.019	-.070
Sense of Well-Being	.712	.206	.202	-.025
Responsibility	.614	.137	.175	.061
Socialization	.531	-.029	-.159	.092
Self Control	.855	-.323	.101	.020
Tolerance	.571	.214	.580	.030
Good Impression	.769	.034	-.019	.056
Communality	.027	.063	-.143	.019
Achievement via Conformance	.705	.305	.135	.058
Achievement via Independence	.214	.017	.781	.034
Intellectual Efficiency	.434	.384	.411	.072

Table 7 Continued

Scale	Factor 1	2	3	4
Psychological Mindedness	.290	.123	.380	-.025
Flexibility	-.125	.0002	.682	.078
Femininity	.036	-.214	.010	.066
Achievement	.063	.151	.084	-.099
Deference	.188	-.139	-.054	-.014
Order	.280	-.292	-.174	-.149
Exhibition	-.156	.302	.030	-.118
Autonomy	-.131	-.061	.075	-.297
Affiliation	.112	.048	.039	.598
Intraception	.121	.139	.081	-.068
Succorance	-.192	.186	.036	.354
Dominance	.229	.427	-.021	-.248
Abasement	-.138	-.311	-.192	.035
Nurturance	.047	-.047	.032	.677
Change	-.023	.065	.128	-.026
Endurance	.155	-.145	-.072	-.101
Aggression	-.227	.047	-.099	-.393
Consistency	.044	.043	.017	-.180
% of Total Variance	33.6	23.3	13.9	10.3
Cumulative Percentage	33.6	56.9	70.8	81.1
Eigen Value	5.7	3.9	2.4	1.7

Study B: Predictive Validity

Question 2

To what extent can a clinician predict a candidate's "success" in real estate sales using the A.I.?

Question 3

To what extent can a clinician predict a candidate's "success" in real estate sales using 10 psychological tests and an interview?

Question 4

To what extent can scales of the A.I. predict a candidate's "success" in real estate sales?

Question 5

To what extent can scales of the E.P.P.S. predict a candidate's "success" in real estate sales?

Question 6

To what extent can scales of the C.P.I. predict a candidate's "success" in real estate sales?

Question 7

To what extent can scales of the A.I., E.P.P.S. and C.P.I. together predict a candidate's "success" in real estate sales?

Question 8

Which methods described in question 2 through 7 best predict sales volume in real estate sales?

To begin to answer questions 2 to 8, a number of correlations and multiple regressions were calculated. These included:

1. Correlations between the criterion and the four A.I. scales, the 15 E.P.P.S. scales and the 18 C.P.I. scales.
2. A correlation between the criterion and a clinician's prediction of success based on the A.I. (see Appendix 3 for results).
3. A correlation between the criterion and a clinician's prediction of success based on 10 psychological tests plus interview (see Appendix 3 for results).
4. A multiple regression among the criterion and the four scales of the A.I..
5. A multiple regression among the criterion and the 15 scales of the E.P.P.S..
6. A multiple regression among the criterion and the 18 scales of the C.P.I..
7. A multiple regression among the three best predictors of each of the C.P.I., E.P.P.S. and A.I..

Table 8 shows the correlations between the four A.I. scales, the 15 E.P.P.S. scales and the 18 C.P.I. scales and the criterion. Reference to Table 8 will indicate that correlations between "success" in real estate sales and the 37 scales of the A.I., E.P.P.S. and C.P.I. range between -.68 and .48. The highest absolute correlation between the A.I. scales and the criterion is .38. Five scales from the C.P.I. and two from the E.P.P.S. correlate at this level or above. These data begin to suggest that there exist better predictors of "success" in real estate than the A.I..

Table 8

Correlations Between the Criterion and the A.I., the E.P.P.S. and the C.P.I.

Aptitude Inventory, California Psychological Inventory, and Edwards Personal Preference Schedule Scales	Success in Real Estate Sales
Intelligent Job Performance	.32
Leadership Qualities	.38
Proper Job Attitude	.25
Relations with Others	-.15
Dominance	-.09
Capacity for Status	-.37
Sociability	-.33
Social Presence	-.26
Self acceptance	-.10
Sense of Well-Being	-.27
Responsibility	.28
Socialization	.26
Self control	-.08
Tolerance	-.53
Good impression	-.17
Communality	.42
Achievement via Conformance	-.02
Achievement via independence	-.22
Intellectual efficiency	-.68
Psychological mindedness	-.38

Table 8 Continued

Aptitude Inventory, California Psychological Inventory, and Edwards Personal Preference Schedule Scales	Success in Real Estate Sales
Flexibility	-.48
Femininity	.06
Achievement	-.31
Deference	-.02
Order	.48
Exhibition	.19
Autonomy	.44
Affiliation	-.22
Intraception	-.12
Succorance	-.15
Dominance	.08
Abasement	.14
Nurturance	-.06
Change	.04
Endurance	-.08
Aggression	-.04
Consistency	.01

Table 9 shows the correlations, multiple regression and R^2's resulting in predicting success in real estate sales. R^2 scores allow for comparison of correlations with the multiple regression. From Table 9 it is evident that the R^2 scores in increasing order of predictability are:

Clinicians using 10 psychological tests plus interview	.014.
A.I. significant scale (1)	.145
Clinician using A.I. alone	.306
E.P.P.S. significant scales (2)	.478
C.P.I. significant scales (3)	.559
A.I., E.P.P.S. and C.P.I. significant scales (4)	.746

These R^2 values indicate the relative ability of the conditions posed in Questions 2 through 7 to predict "success" in real estate sales.

To determine the scales that add significantly to the prediction of "success" when multiple regression was used, a two-way Analysis of Variance was completed. Table 10 shows: the scales used step-wise, the degrees of freedom, F value required at the .05 confidence level, the F observed and whether or not the additional scale adds significantly to the prediction of "success." By inspection of Table 10, it is evident that one A.I. scale, two E.P.P.S. scales, two C.P.I. scales and four scales for the A.I., E.P.P.S. and C.P.I. combined, add significantly to this prediction.

Table 9

Correlations, Multiple Regressions and R^2 to Predict Success in Real Estate Sales

Situation	Correlation	Multiple Regression	R^2
Clinician using Aptitude Inventory alone	.553*		.306
Clinician using Sales Appraisal	.12		.014
Aptitude Inventory significant Scale (Leadership Qualities)		.381*	.145
Edwards Personal Preference Schedule significant Scales (Order, Autonomy)		.691*	.478
California Psychological Inventory significant Scales (Intellectual Efficiency, Communality)		.747*	.559
Aptitude Inventory, Edwards Personal Preference Schedule and California Psychological Inventory Significant Scales (Intellectual efficiency, Communality, Order, Exhibition)		.863*	.746

* Significant at the $P \leq .05$ level.

Table 10

Two Way Analysis of Variance
Adding Scales to the Multiple Regression Step Wise

Instrument	Scales	R^2	Degree of Freedom	F @ .05	F Observed	Significant
Aptitude Inventory	Leadership Qualities					
	Intelligent Job Performance	.216	1,21	4.32	1.63	No
Edwards Personal Preference Schedule	Order					
	Autonomy	.478	1,21	4.32	10.4	Yes
Edwards Personal Preference Schedule	Order					
	Autonomy					
	Exhibition	.502	2,20	3.44	.98	No
California Psychological Inventory	Intellectual efficiency					
	Communality	.559	1,21	4.32	4.76	
California Psychological Inventory	Intellectual efficiency					
	Communality					
	Socialization	.622	2,20	3.44	3.16	No
Best Predictors of the Aptitude Inventory, California Psychological Inventory and Edwards Personal Preference Schedule	Intellectual efficiency					
	Communality	.559	1,21	4.32	4.76	Yes
	Intellectual efficiency					
	Communality					
	Order	.716	2,20	3.44	10.45	Yes
	Intellectual efficiency					
	Communality					
	Order					
	Exhibition	.764	3,19	3.13	3.99	Yes

Table 10 Continued

Instrument	Scales	R^2	Degree of Freedom	F @ .05	F Observed	Significant
	Intellectual efficiency					
	Communality					
	Order					
	Exhibition					
	Relations with others	.777	4.18	2.93	.93	No

Discussion

For a sample of 22 first year real estate sales persons, a clinician was able to predict "success" using the A.I. alone better than he could using a battery of 10 psychological tests plus interview, better than an actuary could using the A.I., but not as well as an actuary could using the E.P.P.S., the C.P.I., or a combination of the A.I., E.P.P.S. and C.P.I.. Further, the A.I. was not found to add significantly to the prediction of success when the A.I., E.P.P.S. and C.P.I. were all used.

It follows that when analyzed by either clinical or actuarial methods, the A.I. does not predict "success" in real estate sales as well as do two other widely used psychological instruments.

Question 9

Does the A.I. add to the actuarial prediction of "success" in real estate sales in a multiple regression when results are available for the A.I., E.P.P.S. and C.P.I.?

By further inspection of Table 10 it is evident that the scales that contribute significantly at the .05 level to the multiple regression are the Ie and Cm scales of the C.P.I., and the Ord and Exh scales of the E.P.P.S..

Discussion

For this sample, the A.I. does not add significantly to the prediction of success in real estate sales.

Additional Findings Pertinent to Study B

As can be seen in Table 7, the R^2 between success in real estate sales and for significant predictors of the A.I., C.P.I. and E.P.P.S. is .746. The regression equation derived from the four significant

predictors is: $Y' = -22.919 + (-1.434\ [Ie]) + (1.960\ [Cm]) + (.416\ [Ord]) + (.235\ [End])$. This equation can be used to predict production records for the first year real estate salesmen. However, because of the small N the 95% confidence interval is $\pm$ 20.64. That is, there is a 95 percent chance that a candidate who scores a Y' of 50 will have a first year sales production of between $29,360 and $70,640. To obtain a smaller standard error, it is recommended that this study be repeated with a larger N.

CHAPTER 5

DISCUSSION AND IMPLICATIONS

In this chapter, results for each study will be discussed in turn, and an attempt to explain significant results will follow. Implications of the findings are presented, and the chapter concludes with suggestions for further research.

Study A: Construct Validity

Study A was designed to determine if relationships among scales of the A.I., E.P.P.S. and C.P.I. are in the direction, and of the magnitude expected by two "experts," and supported by the literature. Table 6 in Chapter 4 indicates that though all but three of the expected correlations are positive (in the expected direction), the correlations are relatively low, between -.07 and .32. These values are lower than those found by Dunnette et al (1958) in correlating the C.P.I. and the E.P.P.S..

These findings indicate that for recruitment and appraisal of male candidates the A.I. does not measure the same dimensions as do the E.P.P.S. and C.P.I.. The findings further suggest that although the A.I. provides definitions of its scales that are highly similar to scales of the more widely validated C.P.I. and E.P.P.S., it does not in fact measure these same dimensions.

Implications

The findings of this study demonstrate the need for more research on the construct validity of the Aptitude Inventory. Before the instrument can be used with confidence, evidence must be furnished of the validity of the four A.I. constructs or scales:

I Intelligent Job Performance

L Leadership Qualities

A Proper Job Attitude

R Relations with Others.

Study B

Study B was designed to assess the predictive validity of the A.I. in predicting "success" in real estate sales. The criterion "success" was established as production records (adjusted for inflation) for the first year in the field of real estate sales. For comparative purposes, two other psychological instruments were also used to predict the criterion; the E.P.P.S. and the C.P.I..

Five R^2 coefficients were compared to determine how the criterion is best predicted. They were developed from:

1. A correlation between a clinician's prediction based on the A.I. and the criterion;
2. A correlation between a clinician's prediction based on 10 psychological tests and interview notes and the criterion;
3. Step-wise multiple regression of the A.I. and the criterion;
4. Step-wise multiple regression of the E.P.P.S. and the criterion;
5. Step-wise multiple regression of the C.P.I. and the criterion;
6. Step-wise multiple regression of the A.I., E.P.P.S. and C.P.I. together and the criterion.

For steps 3 to 6 only scales that added to the multiple regression coefficient at the $p \leq .05$ level were used.

As stated in Chapter 4, the R^2 values in increasing order of ability to predict the criterion were:

Clinician using 10 psychological tests	.014
A.I. significant scale	.145
Clinician using A.I. alone	.306
E.P.P.S. significant scales	.478
C.P.I. significant scales	.559
A.I., E.P.P.S. and C.P.I. significant scales	.746

In the literature, very few studies demonstrate the ability of a clinician to predict better than an actuary. However, in this study the clinician using the A.I. alone predicted "success" in real estate sales better than did an actuary using the same information; R^2 values of .145 and .306.

However, when additional (possibly better) information was available the actuarial prediction improved markedly (R^2 values of .478, .559 and .746), whereas clinical prediction was not as strong, .014. These R^2 values, possibly with the exception of the value obtained for all three psychological tests, are consistent with the literature relating to actuarial prediction. R^2 values of .3 to .6 are common in the predication studies discussed in Chapter 2.

A further important finding was that the A.I. was not found to add significantly to the prediction of the criterion when the A.I., E.P.P.S. and C.P.I. were all used.

It follows from the results of this study that there are better instruments available to predict "success" in real estate than the (Sales) Aptitude Inventory. Two personality inventories that have been found to be better predictors of "success" in this field are

the California Psychological Inventory and the Edwards Personal Preference Schedule.

Implications

The results of this study imply that the A.I. should not be used as a sole predictor in the selection of real estate sales persons. More research on this instrument is required. There are also indications that other personality questionnaires are probably better indicators of "success" in the real estate sales field than the A.I..

The results also imply that it is possible to obtain a valid regression equation to predict first-year sales "success." The equation obtained in this study was:

$Y^1 = -22.919 + (-1.434\ [Ie]) + (1.960\ [Cm]) + (.416\ [Ord]) + (.235\ [End])$.

Though some researchers would say this equation accounts for 74 percent of the variance in predicting success in real estate sales, the low sample size resulted in a large standard errors of measurement, making it of limited value.

A final implication of this study is that before an instrument is used to predict success in an occupation, it must be validated in the setting in which it is being used.

Conclusion

The purpose of this research was to add to the validity information for the Aptitude Inventory. Now, 20 years after the first critical review of the Aptitude Inventory (Siegle, 1959), the same conclusion may be drawn: the A.I. cannot be recommended for any situations but those where extensive research is possible prior to its implementation.

Suggestions for Further Research

1. The most important suggestion for further research is in the area of predictive validity. For the A.I. to be of any use, the ability to predict what it purports to predict must be demonstrated. As it is titled the Employment Aptitude Inventory, the Management Aptitude Inventory, and the Sales Aptitude Inventory, it must be shown that the A.I. can predict success as an employee, a manager, and as a sales person. Studies similar to Study B are recommended to determine the efficacy of the A.I. in predicting success as an employee, a manager, and a sales person.
2. A second important consideration for further research is that related to the validity of the four scales of the A.I.. Recommended research would determine the relationship between supervisors' ratings of employees on the four A.I. scales and results on the A.I.. For such research to be properly conducted, supervisors would have to be very familiar with the subjects, well trained, and validity checks on their ratings would be required.
3. In this study, a regression equation was obtained to predict "success" in real estate sales. However, the sample was small, resulting in a large standard error. Replications of Study B with a larger sample size and using the preferred method of analysis (double-cross validation) are suggested. This would probably produce a regression equation with a small standard error that could be used in screening and selecting real estate sales persons. A sample size of greater than 100 would be required to produce results useful to the industry.
4. Concurrent validity studies are also recommended. To this end,

a replication of Study B could be undertaken. Concurrent validation studies would allow the researcher to get a larger sample as he/she could test a large number of sales persons and get the production records for the past 12 months at the same time. This research would determine the ability of the A.I. to predict "success" in experienced sales persons. Controls for years of experience would be required.

5. Further research into the aspects of A.I. validity not considered in this study are also suggested. Research into Content Validity, Factor Validity, and Congruent Validity are recommended.
6. Further reliability studies of the A.I. are also suggested prior to its continued use.

REFERENCES

Anastasi, A. Psychological testing (3rd ed.). New York: MacMillan Publishing Co., 1968.

Ash, P., & Kroeker, L. P. Personnel selection, classification, and placement. In M. Rosenzweig & L. W. Porter (Eds.), Annual review of psychology. Palo Alto, California: Annual Reviews Inc., 1975, 481-508.

Awang, K. A. A Cross-culture validation study of the vocational preference inventory and the work values inventory, unpublished Phd. thesis.

Bauernfeind, R. H. The matter of "ipsative scores." Personnel and Guidance Journal, 1962, 40, 210-217.

Bray, D. W., & Moses, J. L. Personnel Selection. In P. Mussen & M. Rosenzweig (Eds.), Annual review of psychology. Palo Alto, California: Annual Reviews Inc., 1972.

Broomer, J. Validity information exchange, Personnel Psychology, 1962, 15.

Buros, O. K., (Ed.). Tests in print. Highland Park, N.J.: Gryphon Press, 1974.

Buros, O. K., (Ed.). The fifth mental measurement yearbook. Highland Park, N.J.: Gryphon Press, 1959.

Buros, O. K., (Ed.). The sixth mental measurement yearbook. Highland Park, N.J.: Gryphon Press, 1965.

Buros, O. K., (Ed.). The seventh mental measurement yearbook. Highland Park, N.J.: Gryphon Press, 1972.

Buros, O. K., (Ed.). The eighth mental measurement yearbook. Highland Park, N.J.: Gryphon Press, 1978.

Chronbach, L. J. Essentials of psychological testing (3rd ed.). New York: Harper & Row, 1970.

Chronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1965.

Chronbach, L. J., & Meehl, P. E. Construct validity in psychological tests. In D. N. Jackson & S. Messick (Eds.), Problems in Human Assessment. New York: McGraw Hill, 1967.

Dawes, R. M. The robust beauty of improper linear models in decision making. American Psychologist, 1979, 7, 571-582.

Dawes, R. M., & Corrigan, B. Linear Models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Denton, J. C. A validation of interview type data, Personnel Psychology, 1964, 17, 281-287.

Denton, J. C. Manual, Aptitude inventory. Unpublished Manuscript: Psychological Business Research, 1969.

Dorin, P. A. The use of the California Psychological Inventory in the selection of residence hall staff. Dissertation Abstracts. 1974, 35 (4A), 1906.

Dunnette, M. D. Kirchner, W. K., & DeGedio, J., Relations among scores in Edwards Personal Preference Schedule, California Psychological Inventory, and Strong Vocational Interest Blank for an industrial sample. Journal of Applied Psychology. 1958, 42, 178-181.

Dunnette, M. D. A study of faking behavior on a forced choice self description checklist. Personnel Psychology, 1962, 15, 13-24.

Edwards, A. J. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed.

Journal of applied Psychology, 1953, 37, 90-93.

Edwards, A. L. Manual, Edwards personality inventory. Chicago: Science Research Associates, 1966.

Edwards, A. L. The measurement of personality traits by scales and inventories. New York: Holt, Rinehart and Winston, Inc., 1970.

Edwards, A. L. Edwards personal preference schedule (E.P.P.S.). In D. G. Lake, M. B. Miles, & R. B. Earle, Jr., (Eds.), Measuring Human Behavior. New York: Teachers College Press, 1973.

Goldberg, L. R. Man versus model of man: A rationale, plus some evidence for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-432.

Goldberg, L. R. Five models of clinical judgement: an empirical comparison between linear and nonlinear representations of the human inference process. Organizational Behavior and Human Performance, 1971, 6, 458-479.

Gottfredson, G. D. & Lipstein, D. J. Using personal characteristics to predict parolee and probationer employment stability. Journal of applied Psychology. 1975, 60, 644-648.

Gough, H. G. Clinical vs. statistical prediction in psychology. In L. Postman (Ed.), Psychology in the making. New York: Knopf, 1962.

Gough, H. G. California psychological inventory. In D. G. Lake, M. B. Miles, & R. B. Earle, Jr., (Eds.), Measuring Human Behavior. New York: Teachers College Press, 1973.

Gough, H. G. Manual, California psychological inventory. Palo Alto, Calif.: Consulting Psychologists Press, 1975.

Guion, R. M. Synthetic validity in a small company: a demonstration. Personnel Psychology, 1963, 18, 49-63.

Guion, R. M. Personnel testing. New York: McGraw-Hill, 1965.

Henrichs, J. R. Comparison of "real life" assessments of management potential with situational exercises, pencil-and-paper ability tests, and personality inventories. Journal of Applied Psychology, 1969, 53, 425-432.

Holt, R. R. Clinical and statistical prediction: a reformulation and some new data. Journal of Applied and Social Psychology, 1958, 56, 1-12.

Holt, R. R. Yet another look at clinical and statistical prediction or, is clinical psychology worthwhile? American Psychologist, 1970, 25, 337-349.

Hornaday, J. A., & Aboud, T. A. Characteristics of successful entrepreneurs. Personnel Psychology, 1971, 24, 141-153.

Horstman, P. L. Assessing the California Psychological Inventory for predicting police performance. Dissertation Abstracts. 1977, 37 (12-13), 6387.

Kelly, W. L. Psychological prediction of leadership in nursing. Nursing Research. 1974, 24, 141-153.

Kim, J., & Mueller, C. W. Introduction to Factor analysis. Beverly Hills: Sage Publications, 1978.

Kleinmuntz, B. Personality, measurement, an introduction. Homewood: The Dorsey Press, 1967.

Kuhlen, R. Needs perceived satisfaction opportunities and satisfaction with occupation. Journal of applied psychology. 1963, 47, 56-64.

Lake, D. G., Miles, M. B., & Earle, R. B. Measuring Human Behavior. New York: Teachers College Press, 1973.

Lanyon, R. I., & Goodstein, L. D. Personality Assessment. New York: John Wiley & Sons, Inc., 1971.

Loevinger, J. Objective tests as instruments of psychological theory. In D. N. Jackson & S. Messic (Eds.), Problems in human assessment. New York: McGraw Hill, 1967.

Longstaff, H. P. Fakability of the Strong Interest Blank and the Kuder Preference Record. Journal of Applied Psychology, 1948, 32, 360-369.

Marks, P. A., & Seeman, W. The actuarial description of abnormal personality. Baltimore: The Williams & Wilkins Company, 1963.

McKee, J. P., & Turner, W. S. The relation of "drive" ratings in adolescence to C.P.I. and E.P.P.S. scores in adulthood. Vita Humana, 1961, 4, 1-14.

Meehl, P. F. Clinical versus statistical prediction. Minneapolis: University of Minnesota Press, 1954.

Megargee, E. The California psychological inventory handbook. San Francisco: Jossey-Bass Inc., 1972.

Mischel, W. Personality and assessment. New York: John Wiley and Sons, Inc., 1968.

Muchensky, P. M., & Hoyt, D. P. Predicting vocational performance of engineers from selected vocational interest personality, and scholastic aptitude variables. Journal of Vocational Behavior. 1974, 5, 115-123.

Norrell, G., & Grater, H. Interest awareness as an aspect of self awareness. Journal of Counselling Psychology, 1960, 7, 289-292.

Nunnally, I. C. Educational measurement and evaluation (2nd ed.). New York: McGraw-Hill, 1964.

Nunnally, J. C. Introduction to psychological measurement. New York: McGraw-Hill, 1970.

Nunnally, J. C. Introduction to statistics for psychology and education. New York: McGraw-Hill, 1975.

Ross, P. F. Validity information exchange. Personnel Psychology, 1963, 16, 283-288.

Sawyer, J. Measurement and prediction, clinical and statistical. Psychological Bulletin, 1966, 66, 178-200.

Scissons, E. H. Convergence of clinical judgment: a multivariate analysis. Unpublished Phd. thesis.

Shertzer, B. & Linden, J. Fundamentals of individual appraisal. Boston: Houghton Mifflin Company, 1979.

Siegel, L. Review of the aptitude inventory. Journal of Counselling Psychology, 1959, 6, 319-320.

Simpson, R. H. The specific meanings of certain terms indicating different degrees of frequency. Quarterly Journal of Speech, 1944, 30, 328-330.

Slocum, J. W., & Hand, H. H. Prediction of job success and employee satisfaction for executives and foremen. Training and Development Journal, 1971, 25, 28-36.

Spitzer, M. E., & McNamara, W. J. A managerial selection study. Personnel Psychology, 1964, 17, 19-40.

Tatsuoka, M. M. Validation studies, the use of multiple regression equations. Champaign: The Institute for Personality and Ability Testing, 1969.

Tatsuoka, M. M. Multivariate analysis Techniques for educational and psychological research. New York: John Wiley & Sons, Inc., 1971.

Technical recommendations for psychological tests and diagnostic techniques. Psychological Bulletin Supplement, 1954, 51, Part 2, 1-38.

Wessman, A. G. Faking personality test scores in a simulated employment situation. Journal of Applied Psychology, 1952, 36, 112-113.

Wiggens, J. S. Personality and prediction Principles of personality assessment. Reading: Addison-Wesley Publishing Co., 1973.

Woodworth, R. S. History of psychology in autobiography, Vol. 2. New York: Russell & Russell, 1951.

APPENDIX 1

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I.

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I.

Scale	I	L	A	R	Do	Cs	Sx	Sp	So	Wb	Re	So	Sc
I	1.0	.09	-.01	-.19	0.10	.11	.10	.01	.01	.19	.17	.16	.24
L		1.0	.08	-.12	.29	.20	.23	.18	.19	.13	.05	-.01	.02
A			1.0	-.23	.06	-.05	.06	.04	.07	-.01	-.07	.00	-.10
R				1.0	-.04	-.03	-.03	-.04	-.04	-.01	.06	.03	.03
Do					1.0	.49	.56	.39	.48	.23	.36	.06	-.08
Cs						1.0	.56	.47	.36	.35	.29	.05	.06
Sx							1.0	.59	.56	.33	.19	.11	-.08
Sp								1.0	.50	.21	-.09	-.17	-.30
So									1.0	-.01	-.07	-.11	-.42
Wb										1.0	.50	.31	.57
Re											1.0	.40	.50
So												1.0	.44
Sc													1.0
To													
Gi													

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I. Continued

Scale	To	Gi	Cm	Ac	Ai	Ie	Py	Fx	Fe	Ach	Def	Ord	Exh
I	.17	.24	.02	.26	.11	.17	.14	-.08	-.02	.17	.03	.23	-.06
L	.06	.18	-.08	.17	-.05	.07	.07	-.07	-.07	.16	.01	-.06	.13
A	-.12	-.01	.01	-.05	-.18	-.09	-.03	-.15	-.04	-.06	.02	.06	.04
R	.07	.01	.03	-.00	.05	-.01	-.05	.09	.07	-.22	.06	-.07	-.09
Do	.26	.15	.04	.36	.08	.33	.16	-.06	-.10	.16	-.08	-.17	.18
Cs	.44	.27	-.07	.38	.29	.46	.26	.17	-.10	.18	-.02	-.17	.14
Sy	.31	.21	.07	.38	.07	.43	.18	.01	-.18	.13	-.11	-.16	.25
Sp	.29	-.06	-.07	.16	.20	.43	.25	.33	-.31	.14	-.26	-.34	.33
Sa	.01	-.17	.10	.08	.01	.16	-.02	-.01	-.09	-.10	-.20	-.24	.25
Wb	.58	.59	-.01	.58	.29	.49	.31	.09	-.10	.12	.09	.09	-.02
Re	.50	.45	.11	.54	.32	.36	.23	-.05	.16	.08	.17	.14	-.15
So	.17	.37	.17	.40	.01	.17	.06	-.22	.14	.03	.16	.23	-.10
Sc	.47	.75	-.03	.54	.28	.24	.27	-.02	.17	.02	.28	.34	-.27
To	1.0	.45	-.05	.52	.61	.57	.40	.31	.02	.10	.05	-.01	-.03
Gi		1.0	-.14	.56	.17	.26	.25	-.06	.06	.09	.24	.24	-.13
Cm			1.0	.05	-.08	-.06	-.14	-.22	.09	-.04	.08	.12	-.06

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I. Continued

Scale	To	Gi	Cm	Ac	Ai	Ie	Py	Fx	Fe	Ach	Def	Ord	Exh
Ac				1.0	.28	.52	.28	-.02	.04	.18	.14	.14	-.05
Ai					1.0	.38	.33	.56	.10	.11	-.03	.12	-.01
Ie						1.0	.33	.21	-.17	.17	-.03	-.09	.04
Py							1.0	.27	-.06	.14	-.01	-.05	-.01
Fx								1.0	-.01	.03	-.10	-.31	-.06
Fe									1.0	-.06	.13	.11	-.14
Ach										1.0	.02	-.09	.11
Def											1.0	.29	-.23
Ord												1.0	.26
Exh													1.0
Aut													
Aff													
Int													
Suc													
Dom													
Aba													

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I. Continued

Scale	Aut	Aff	Int	Suc	Dom	Aba	Nur	Chg	End	Agg	Con
I	-.06	-.08	.02	-.08	.11	-.18	-.06	-.07	.17	-.02	.09
L	.06	-.12	.08	-.18	.32	-.16	-.12	-.01	.02	.05	.11
A	-.09	-.11	-.02	-.06	.06	.04	-.10	-.05	.21	.05	.05
R	-.11	.26	.13	.11	-.17	.13	.27	.01	-.15	-.17	-.17
Do	-.05	.03	.10	.07	.46	-.25	-.01	-.02	-.07	.02	.07
Cs	-.01	.09	.12	-.15	.27	-.34	-.02	.11	-.09	-.09	.05
Sy	-.09	.08	.07	-.15	.34	-.32	-.02	.10	-.05	-.03	.03
Sp	.09	.04	.01	-.10	.25	-.39	-.09	.22	-.18	.07	.08
Sa	.03	-.02	.04	-.09	.27	.24	-.10	.07	-.15	.12	.03
Wb	-.07	.09	.14	-.21	.24	.23	-.01	.07	.10	-.18	.06
Re	-.18	.10	.15	-.08	.19	-.05	.13	-.11	.09	-.22	.03
So	-.21	.10	.02	-.01	.06	-.01	.07	-.10	.13	-.14	.01
Sc	-.14	.07	.13	-.11	.08	-.02	.07	-.06	.22	-.28	.02
To	-.04	.14	.14	-.13	.19	-.25	.04	.09	-.01	-.21	.04
Gi	-.16	.08	.20	.23	.20	-.11	.05	-.03	.19	-.27	.03
Cm	-.19	.02	-.04	.08	-.03	.07	.06	-.09	.06	.02	-.01

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I. Continued

Scale	Aut	Aff	Int	Suc	Dom	Aba	Nur	Chg	End	Agg	Con
Ac	-.19	.08	.13	-.15	.33	.24	.03	-.04	.10	-.21	.06
Ai	.06	.11	.06	.03	.07	-.27	.03	.11	-.13	-.14	.03
Ie	-.03	.10	.11	-.09	.22	-.32	.10	.11	.03	-.14	.06
Py	.05	.01	.08	-.12	.16	.21	-.04	.07	.03	-.07	.03
Fx	.14	.09	.04	.05	-.02	-.15	.05	.21	-.22	-.06	-.01
Fe	-.11	.03	.06	.13	-.14	.14	.10	-.09	.04	-.11	-.01
Ach	.09	-.16	-.07	-.13	.20	-.34	-.26	.01	.08	-.04	.24
Def	-.31	-.05	.13	-.07	-.04	.09	-.03	-.21	.14	-.22	-.08
Ord	-.20	-.20	-.06	-.12	-.04	.07	-.11	-.24	.38	-.14	-.01
Exh	.11	-.09	-.15	-.05	.10	-.22	-.24	.09	-.23	.14	.08
Aut	1.0	-.17	-.15	-.14	.05	-.15	-.26	.18	-.24	.20	.09
Aff		1.0	-.03	.16	-.16	-.09	.40	.04	-.23	-.29	-.16
Int			1.0	.22	.01	.05	.01	-.08	-.05	-.16	-.08
Suc				1.0	-.26	.02	.20	-.11	-.20	-.06	-.10
Dom					1.0	-.27	-.22	-.07	-.04	.10	.16
Aba						1.0	.15	-.26	.07	-.03	-.14

Correlations Among Scales of the A.I., E.P.P.S. and C.P.I. Continued

Scale	Aut	Aff	Int	Suc	Dom	Aba	Nur	Chg	End	Agg	Con
Nur							1.0	-.16	-.09	-.28	-.18
Chg								1.0	-.22	-.08	.04
End									1.0	-.09	.07
Agg										1.0	.13
Con											1.0

APPENDIX 2

Clinician's Prediction of "Success"
Based on Aptitude Inventory alone, and
Ten Test Battery

Clinician's Prediction of "Success" Based on Aptitude Inventory alone, and Ten Psychological Test Battery

Subject	Intelligent Job Performance	Leadership Qualities	Proper Job Attitude	Relations With Others	Rating Based on Aptitude Inventory Results	Rating Based on Ten Psychological Test Battery
1	90	80	60	40	1	4
2	20	01	05	20	5	5
3	80	70	80	70	1	2
4	30	01	80	95	3	2
5	80	90	60	10	2	1
6	50	60	70	80	1	3
7	05	60	60	30	4	4
8	30	70	80	90	2	1
9	10	30	95	30	4	2
10	60	70	30	80	4	2
11	70	30	90	50	1	1
12	70	30	20	60	4	5
13	20	05	20	80	5	5

Clinician's Prediction of "Success" ... Continued

Subject	Intelligent Job Performance	Leadership Qualities	Proper Job Attitude	Relations With Others	Rating Based on Aptitude Inventory Results	Rating Based on Ten Psychological Test Battery
14	60	30	20	90	4	4
15	101	90	05	80	5	5
16	40	05	70	10	4	1
17	80	01	90	40	2	2
18	05	10	01	60	5	4
19	40	90	10	95	4	3
20	40	05	05	70	5	5
21	01	80	90	30	3	2
22	10	01	05	90	5	4